

## Comments from JABA Readers

### ANOTHER'S VIEW OF OBSERVER AGREEMENT AND OBSERVER ACCURACY

There have been several issues related to research practices in applied behavior analysis that have received considerable attention in the literature. Not least among these is the topic of observer agreement. Since Powell, Martindale, and Kulp (1975) explained how certain methods of observer agreement are unsuitable for assessing some kinds of observational practices, many others have reported research clarifying and extending their findings.

The topic that has developed into the most important concern in these discussions, however, is the question of whether or not observer agreement, a necessary component for publishing a study using observational data, is even a relevant issue. The argument has been made that the critical concern is, or should be, the accuracy of measurement, not the agreement (or reliability) of observation (see Johnston & Pennypacker, 1980).

The stress on agreement rather than on accuracy implies the ability of reliable, multiple observers to determine the exact status or amount of some event or behavior. It is not uncommon for reliability and accuracy to coincide; neither is it uncommon for there to be little or no relationship between the two. Even when two observers are in total agreement, there is no necessary relation between that agreement and the actual occurrence of the observed event or action. Both observers could be equally wrong.

There are occasional reports in the press about independent observers reporting the landing of a spacecraft and describing their subsequent (independent but remarkably similar) flights to Venus. We rarely have complete faith in such reports; rather, we look for some other evidence concerning

the accuracy or inaccuracy of the reports. They may have seen the same TV show or movie; they may even be members of the same intergalactic organization. The same logic should hold for the reliable reports of two observers on talking-in-class or self-mutilation. These observers may have been (mis)trained by the same experimenter or they may have discussed a strategy that would increase their reliability, even at the expense of accuracy.

Applied behavior analysts are aware that editorial practices have not changed as a result of these arguments or the earlier data on other issues of observer agreement. A comment from a different source on a different topic might clarify a part of this argument and help behavior analysts better evaluate current editorial practices. The philosopher Wittgenstein was also concerned with the accuracy of statements about various matters. In commenting on problems encountered when one is unable to test or confirm some event for correctness, he used the following analogy, which I find relevant to the behavioral discussions on agreement and accuracy: "As if someone were to buy several copies of the morning paper to assure himself that what it said was true" (Wittgenstein, 1953, p. 94).

### REFERENCES

- Johnston, J. M., & Pennypacker, H. S. (1980). *Strategies and tactics of human behavioral research*. Hillsdale, NJ: Erlbaum.
- Powell, J., Martindale, A., & Kulp, S. (1975). An evaluation of time-sample measures of behavior. *Journal of Applied Behavior Analysis*, 8, 463-469.
- Wittgenstein, L. (1953). *Philosophical investigations*. New York: Macmillan Publishing Co., Inc.

SAMUEL M. DEITZ  
Georgia State University

Received July 21, 1987

Final acceptance September 23, 1987